# High Performance
# Human-Computer Interfaces





JASON
MITRE

# High Performance Human-Computer Interfaces

Study Leaders:
A. Despain
R. Westervelt

Contributors Include:

H. Abarbanel
S. Block
C. Callan
J. Cornwall
W. Dally
R. Garwin
R. Henderson
S. Koonin
W. Press
J. Vesecky
E. Williams
H. Woodin

JSR-96-130

September 1997



JASON
The MITRE Corporation
1820 Dolley Madison Boulevard
McLean, Virginia 22102-3481
(703) 883-6997

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



**1. AGENCY USE ONLY (Leave blank)**

**2. REPORT DATE**
September 25, 1997

**3. REPORT TYPE AND DATES COVERED**

**4. TITLE AND SUBTITLE**
High Performance Human-Computer Interfaces

**5. FUNDING NUMBERS**
13-988534-04

**6. AUTHOR(S)**
A. Despain, R. Westervelt, H. Abarbanel, S. Block,d C. Callan, J. Cornwall, W. Dally, R. Garwin, R. Henderson, S. Koonin, W. Press, J. Vesecky, E. Williams, H. Woodin

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
The MITRE Corporation
JASON Program Office
1820 Dolley Madison Blvd
McLean, Virginia 22102

**8. PERFORMING ORGANIZATION REPORT NUMBER**
JSR-96-130

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
Advanced Research Projects Agency
3701 North Fairfax Drive
Arlington, Va. 22203-1714

**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**
JSR-96-130

**11. SUPPLEMENTARY NOTES**

**12a. DISTRIBUTION/AVAILABILITY STATEMENT**
Approved for public release; distribution unlimited.

**12b. DISTRIBUTION CODE**
Distribution Statement A

**13. ABSTRACT** *(Maximum 200 words)*

Human interfaces to the computer have remained fairly crude since the use of teletypes despite the fact that computer, storage and communication performance have continued to improve by many orders of magnitude.

How much better can we do? We will try to show in this presentation that an improvement of more than 10 times in effective bandwidth should be achievable.

**14. SUBJECT TERMS**

**15. NUMBER OF PAGES**

**16. PRICE CODE**

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| Unclassified | Unclassified | Unclassified | SAR |

# Briefings and Contributions

M. Arbib, USC
J. Grossman, NRAD
E. Hovy, ISI
D. Karron, NYU
H. Kolb, Univ. of Utah
R. Normann, Univ. of Utah
G. Ojemann, Univ. of Washington

T. Sejnowski, Salk
T. Starner, MIT
D. Tank, Bell Labs
A. Van Dam, Brown Univ.
M. Weiser, Xerox
J. Wyatt, MIT

# Briefings and Contributions

We are greatly indebted to these individuals who briefed us on various aspects of the Human-Computer Interface. Their materials and discussions have made this study possible.

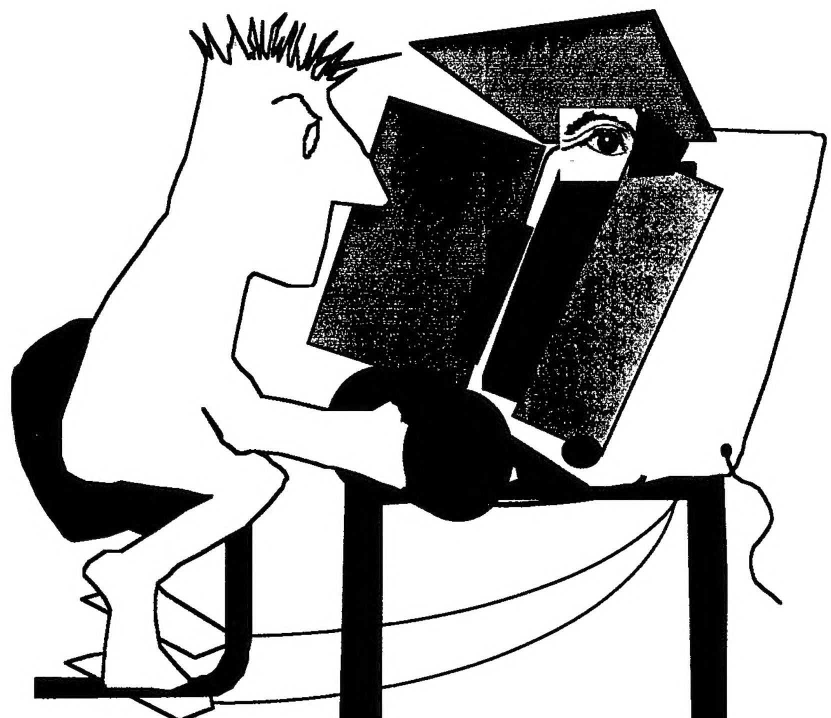
# Interfaces: Can We Do Better?

# Interfaces: Can We Do Better

Human interfaces to the computer have remained fairly crude since the use of teletypes despite the fact that computer, storage and communication performance have continued to improve by many orders of magnitude.

How much better can we do? We will try to show in this presentation that an improvement of more than 10 times in effective bandwidth should be achievable.

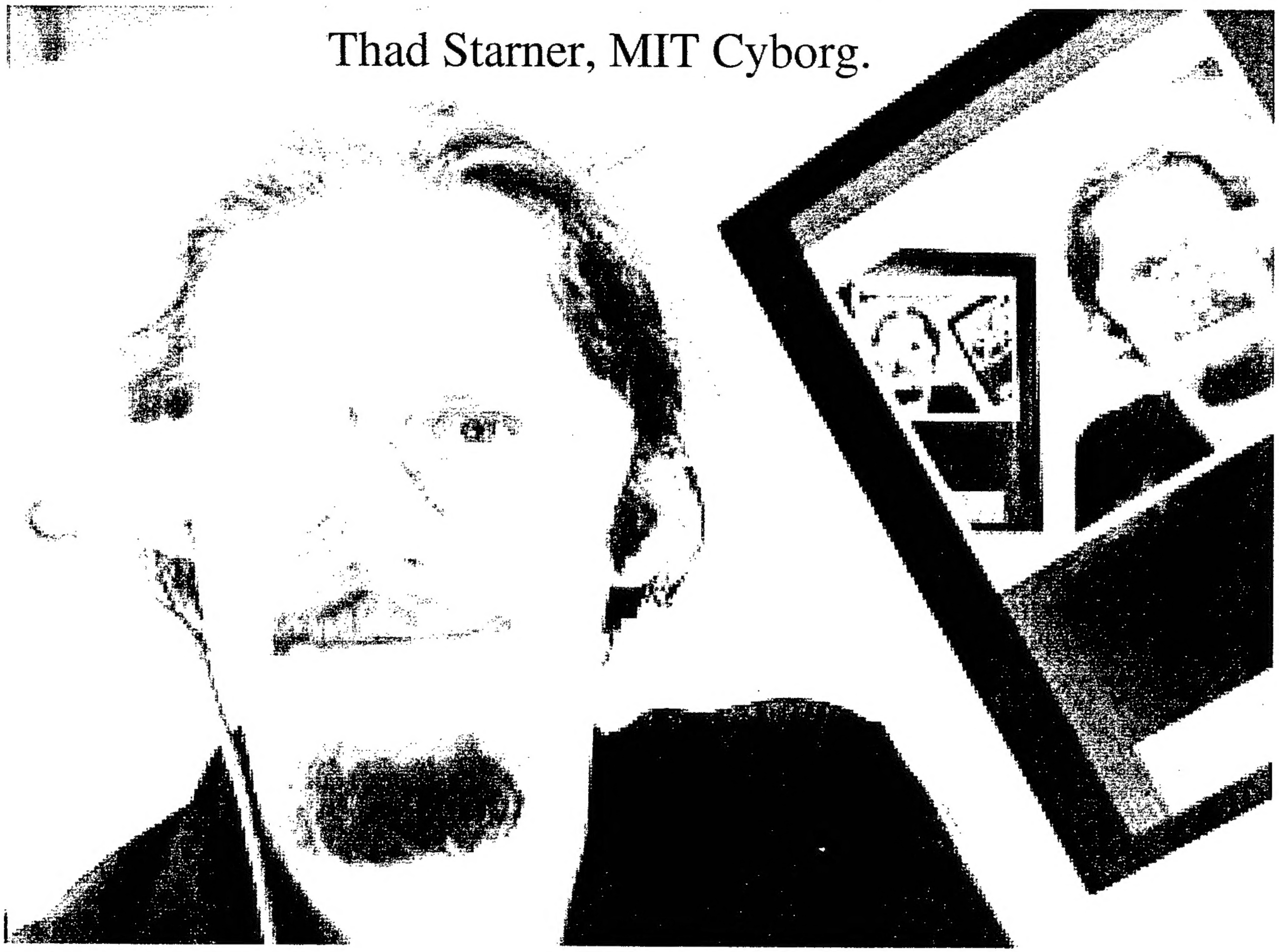
# An Answer for the Office ?



# An Answer for the Office?

Looking ahead to the conclusions, it appears that an unobstrusive, high-performance interface can be obtained by “gussying-up” the familiar professional workstation.

# An Answer for the Field ?


Thad Starner, MIT Cyborg.

# An Answer for the Field?

This is a photo of Thad Starner, who briefed us on his cyborg experiment on himself. His report on this experience as a cyborg had a large influence on our thinking. Thad has adapted relatively inexpensive COTS components into a very effective, if intrusive interface. Again it appears, as we shall see, that the intimate cyborg interface can be greatly improved and can serve professionals involved in critical field tasks, such as fighting wars, floods and fires.

# Overview

- Introduction
- Direct Neuron Connections
- Haptic (Direct Contact) Interfaces
- Feedback-loop Interfaces
- User Interfaces
- Digital Personal Communication Assistants
- Professional Workstation
- Cyborg System
- Conclusions and Recommendations

# Overview

We will first introduce the context of the study and some important general concepts. Next we will examine possibilities for directly wiring-up computers to humans via neural implants. Haptic (direct contact) interfaces such as keyboards and data gloves are considered next. Unobtrusive (non-direct contact) interfaces such as video cameras observing the human are then considered. The critical feature to achieve high performance with such interfaces is as a part of a feedback sensor-effector loop. User interfaces and the concept of a digital personal communication assistant can also improve interface effectiveness. The conclusions concern workstations, cyborgs, techniques and some possible projects.

# Introduction

- Work Statement
- Context -- Study Terms of Reference
- Role of Training and Learning
- Channel Bandwidth vs Knowledge Base
- Levels of Abstraction

# Introduction

In the introduction section, we provide the workstatement, essentially the goals set out for us by DARPA. We then discuss the terms of reference for the study. The study is directed towards improving the interface for professionals that are willing to undergo extensive training and learning. We conclude the introduction with several critical, if general concepts employed in the study.

# Work Statement

JASON will examine and develop new concepts that show promise in providing high performance in human-computer interfaces (HPHCI .)

The performance of computers, data storage and network systems is rapidly increasing each year. To the contrary, the human interface to these systems is only improving slowly, if at all. How can higher performance human-computer interfaces be achieved ?

JASON will take a long-term view and examine a wide range of both hardware and software possibilities:

1. Direct connections to neurons,
2. Better "impedance-match" to human sensors and effectors,
3. Intelligent user interfaces.

# Work Statement

Howard Frank and Allen Sears of DARPA asked JASON to consider what concepts might help provide a higher performance human-computer interface. This task is described by this work statement.

# Context

- Long term view: 5-10 years
- Study is not a survey or program review
- Study offers suggestions for future research
- Focus is on interfaces for professionals:
  - Intelligence analysts, military planners, managers
  - Scientists, engineers, designers, technicians
  - Teachers, librarians, researchers
  - Operators: pilots, etc.
  - Elite teams: seals, rangers, firefighters
- Focus is not on interfaces for general public as needed for games, entertainment, sales, training

# Context

The terms of reference for our study were to take a long-term view and concentrate on new ideas that might inspire future research activity. We chose to focus on interfaces for professionals rather than interfaces for the mass market: games, entertainment, sales and training.

# Context

(concluded)

- Goal is to improve effective interaction bandwidth, even if this requires extensive training
- Standards and COTS are less important than:
  - Special purpose solutions
  - Customization
  - Adaptability

# Context

(concluded)

For mass market applications, standards and ease of use are the most important issues. This is not true for professional computer users who presumably are willing to learn new interface protocols and be trained in their most effective use in order to greatly increase their performance in interfacing with their computers.

# Role of Training and Learning

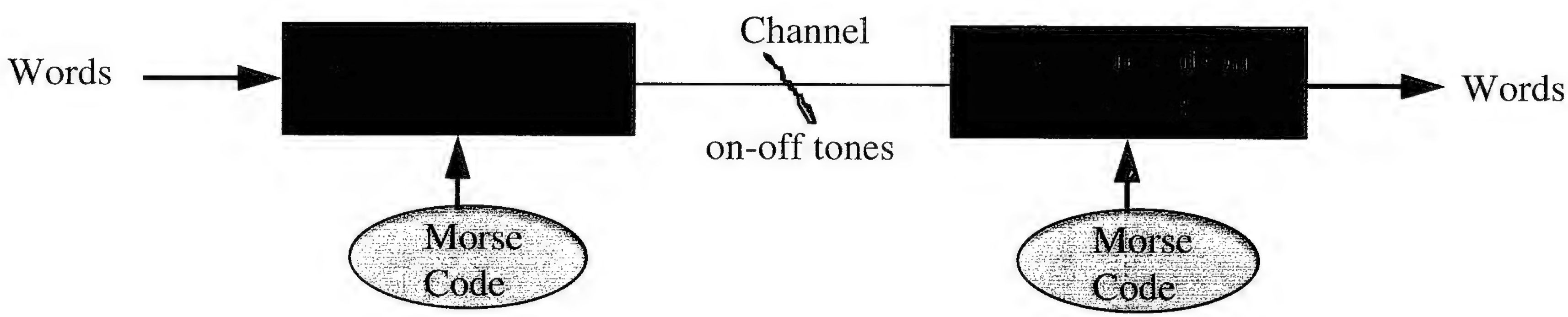
- Commercial development of user interfaces aims for greatest ease of use for average user
  - Ease of use : ↑
  - Training required: ↓
  - Performance (bandwidth) ?
- Professional user interfaces can require training with the reward of increased performance
  - Good typist : ~80 wpm = ~53bps
  - Shorthand : ~120 wpm = ~80bps
  - Court stenographer : ~200wpm = ~133bps
  - Concert pianist : ~300 bps

# Role of Training and Learning

Professional users of machine interfaces can, with extensive training, achieve large improvements over casual users.

# Shared Knowledge

- Communication across any interface channel is impossible without shared knowledge. Example:
  - Given a single stimulus such as a single tone that can be turned off and on, then the sender and receiver can only communicate if they both share an encoding, e.g. Morse code.



# Shared Knowledge

Consider a very simple interface that consists of a single simple tone that can only be switched off and on.  To employ this interface at all, there must be common (shared) knowledge on each side of the interface. An example is the Morse code system in which each side of the interface has a copy of the Morse code.  If both sides don't share something equivalent to the Morse code, they can't communicate at all.

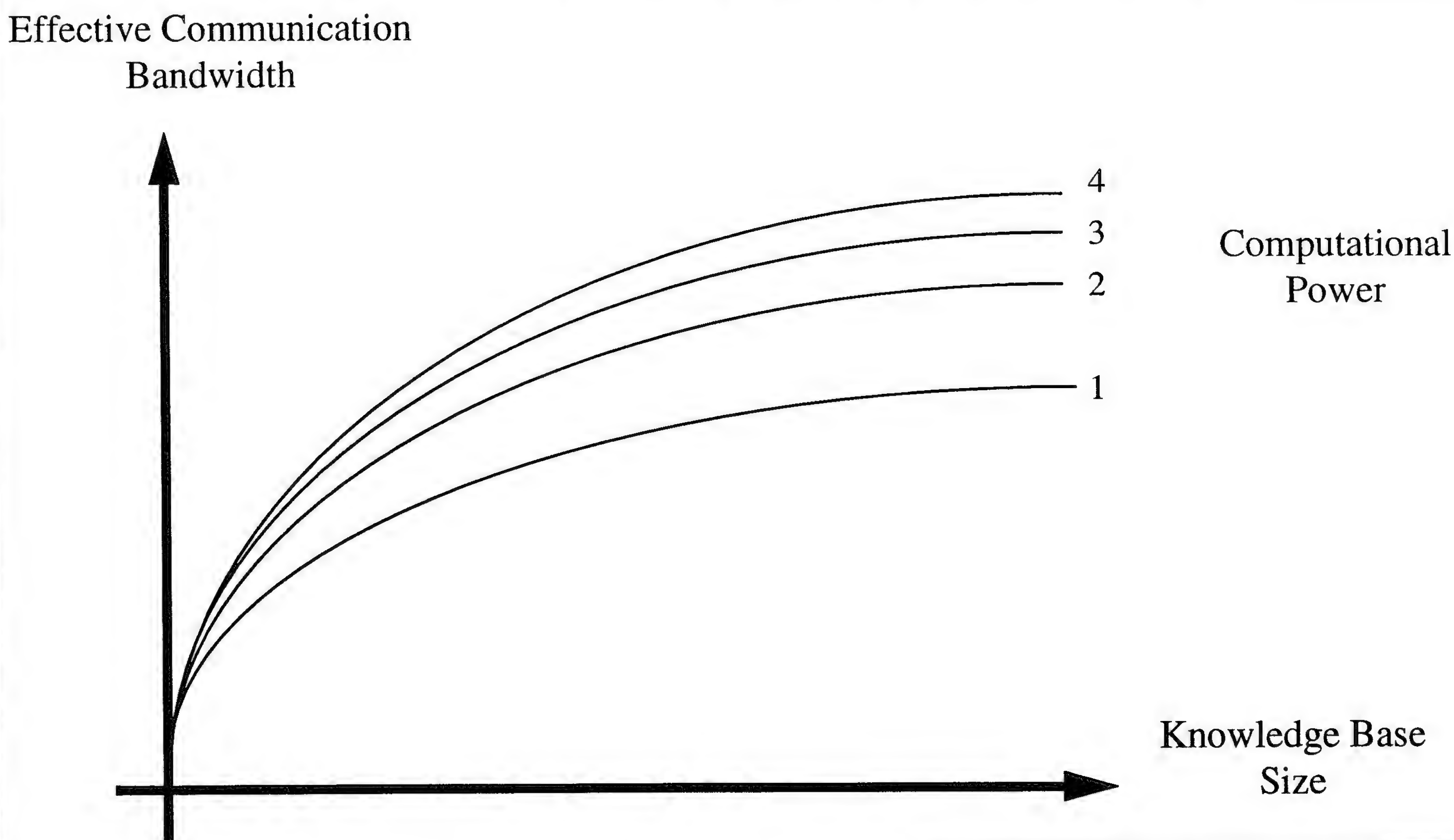
# Channel Bandwidth vs Knowledge Base

- The effective bandwidth through a channel depends how much knowledge is shared by each entity attempting to communicate.
- Such knowledge can be static within a communication session.
  - Example: ‘Twenty questions’ where the players share nearly identical knowledge & experience.
- Dynamic knowledge can also be constructed during a session.
  - Example: Dynamic data compression such as the GZIP utility [Lempel-Zif Algorithm].
- All abstraction levels will need a knowledge base.

# Channel Bandwidth vs Knowledge Base

The effective bandwidth across an interface is a function of how much knowledge is shared by the entities on each side of the interface. Such knowledge can be static such as our previous Morse code example or dynamic such as the cache stack build-up by the Lumpel-Zif data compression algorithm.

# Tradeoffs: Computation, Knowledge and Communication Bandwidth



# Tradeoffs: Computation, Knowledge and Communication Bandwidth

There is also a well-known trade-off between computation and quantity of knowledge. As a result, in general, more computation can be employed to improve effective interface bandwidth.

# Knowledge Properties

- Quantity: The larger the shared knowledge base, the more effectively the channel can be used.
- Static: Training is a process of acquiring static knowledge that can increase effective bandwidth.
- Dynamic: Interface adaptation jointly increases the stored knowledge at the sender & receiver.
- Abstraction: Knowledge, of different forms, must exist at all levels of the communication hierarchy.

# Knowledge Properties

The type of knowledge needed at each level of abstraction of the interface hierarchy will be very different level to level. However, each abstract interface will have an identifiable grammar and knowledge representation method, generally in the form of rules.

For humans to acquire skill in the use of the abstract interface, the grammar and knowledge of the interface must be learned through training. Adaption, by learning knowledge during an interface session will be especially important for the computer system.

# Asymmetric Channels

- Effective bandwidth in one direction through a channel can be increased by employing the reverse channel. Example:
  - The game of Twenty Questions
    - forward channel: 20 bits (the twenty yes or no answers)
    - reverse channel: About twenty sentences or ~20 k bits
    - ratio is ~1000:1
    - goal is an answer of about ~1000 bits
  - Note that Twenty Questions can only be played between two people who share significant knowledge and/or experience, perhaps $10^{10}$ bits!
    [ Human genome, language, culture, experience]

# Asymmetric Channels

The direct human interface is asymmetric in the sense that vision provides far more bits into a human than the human effector system can provide output. It is interesting that excess bandwidth in one direction can be employed to increase the effective bandwidth in the other direction.

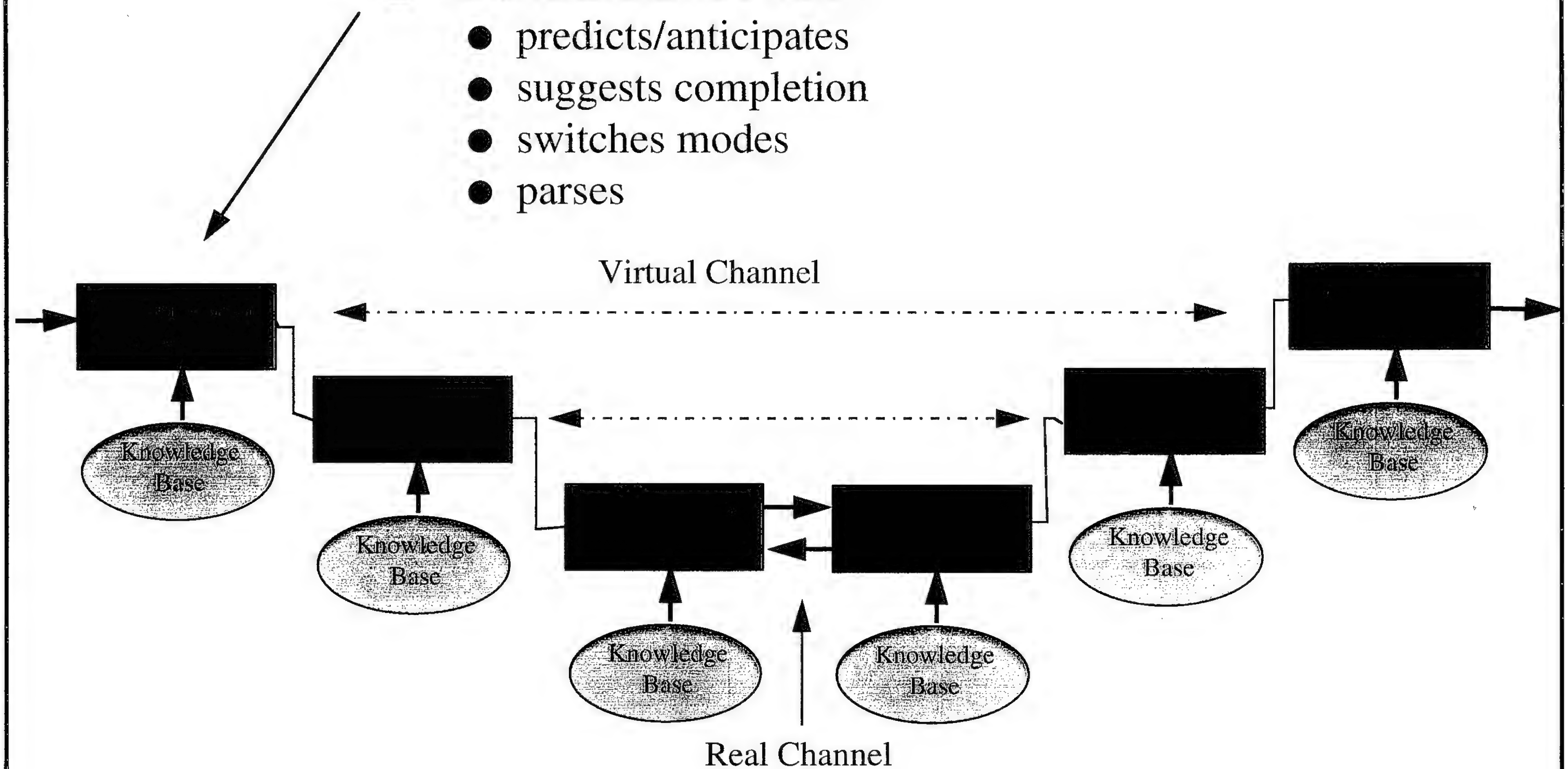
A good example is the child's game of "twenty questions". The first player thinks of any definite object whatsoever. He then answers up to twenty questions put to him by the second player with only a yes or no answer. The first player wins if the second player cannot completely describe the object within the twenty question limit.

The forward channel (across this human to human interface) is composed of twenty bits (or less) total. The reverse channel is a total of twenty questions or about 20,000 bits total. If the second player wins, (and this happens about half the time), then about 1000 bits, (the object description) are effectively sent across the forward channel within the 'physical' 20 bits. The ratio of 'improvement' is about 1000:1 due to use of both the wide bandwidth reverse channel and a very large shared knowledge base. Exactly how large the shared knowledge must be is not clear but may be as much as $10^{10}$ bits since possibly it might be necessary to include the entire human language, culture and experience.

# Communication Channel Hierarchy

## [A protocol stack and more]

- Communication Process
  - predicts/anticipates
  - suggests completion
  - switches modes
  - parses



# Communication Channel Hierarchy

## [A protocol stack and more]

The means humans employ to communicate with each other and their machines are very complex, involving layers of protocol, language, grammar and syntax. We will employ a model of this which we call the Communication Channel Hierarchy. It has the form of a formal communication protocol stack and identifies both a knowledge base and communication process at each level.

# Data Compression

- One-way data flow compressed by:
  - Use of reverse channel
  - Use of knowledge base
  - Change of data coding (static and dynamic)
  - Error coding and error correction

# Data Compression

Classic data compression techniques also employ such communication channel hierarchies in which reverse channels, shared knowledge bases and error detection and correction are all employed.

# Channel Language Grammar

Push Left Button

- Grammar makes a difference in performance even to an expert.
- Experts learn big chunks of grammar and no longer think about the details ....
- Yet other encodings of the chunks could result in higher performance communication.

Push Right Button

[redacted]

- A direct text move command would be something additional to learn but would result in more efficient communication.

Release Buttons Here

# Channel Language Grammar

The grammar chosen to define the language employed by a virtual channel can have a profound effect on the effective bandwidth across the interface. Novices, in using a given channel need a small but complete set of simple expressions to employ. Experts on the other hand can learn a larger vocabulary with additional expressions that represent a series of the simpler expressions. Often, however, they do not because they internalize big chunks of expressions into a single higher level expression in their own thinking. As a result while they think more efficiently, they continue to employ a low-performance grammar at the human-computer interface that has become automatic to use.

A small amount of training of such experts can lead to large improvements. Consider the Apple Mac graphical user interface (GUI). It is elegant and simple, easily employed by anyone from novices of four years of age up to experts. It is possible to keep all the simple expressions of the Mac GUI but add more complex ones for experts to use. For example the task of moving objects (text, etc.) in a document requires a number of simple commands:
1) Select start of text; 2) Select end of text; 3) Pull down edit menu; 4) Select 'cut'; 5) Select position to place text; 6) Pull down edit menu; 7) Select 'paste.' All these actions can be compressed into a single complex action involving a second button of the mouse: 1) Select start of text with a right button push and hold; 2) Select end of text with left button push and hold; 3) Move to location to place text and release the buttons, Note the more than two to one speed-up in this example.

# Levels of Abstraction - Connections

- [0]: Neural interface (Retina,Cortex,Motor nerves)
- [1]: Human sensor stimulation and movement
- [2]: Symbol recognition and stroke generation
- [3]: User Interfaces (Text,Windows,Icons,Graphics)
- [4]: Digital Personal Communication Assistant
- [5]: Networked Systems
  - Groups / Data Bases / Nets
  - [Web-based Interfaces into 'Digital Societies']
- [-]: Integrated [ Wearable computers & Cyborgs]

# Levels of Abstraction - Connections

Some major levels of abstraction of the communication channel hierarchy are identified in this illustration. We will, in the pages to follow, discuss techniques to improve the human-computer interface at each of these levels.

# Direct Neuron Connections

- Direct connections to neurons can be made -- Today these remain functional for :
  - Cochlear implants (human) = years
  - Retina (rabbit) = days
  - Visual cortex (cats) = 9 months
  - Motor nerves (humans) = years
- There is active research to implement such implants to aid the deaf and blind.
- For direct connections, the visual cortex offers the greatest potential for high bandwidth

# Direct Neuron Connections

On the human side of the interface, the ultimate physical source and sink of all communications are the neurons of the brain. The question then arises: "Is it possible to achieve very high performance through the human-computer interface by directly wiring to neurons?" Today John Wyatt at MIT successfully implants electrodes into the neurons of the retina. Similarly Richard Normann at Utah and others directly connect to neurons in the visual cortex. These connections last for up to nine months before they fail. Cochlear and motor nerve implants apparently last for years. The object of these implants is to bring some sight to the blind, some hearing to the deaf and some control to disconnected muscles.

# Sensing Neuronal Activity

- EEG: Electric fields generated by neurons are integrated, distorted and attenuated by tissue. Only gross electrical behavior can be observed
- MEG: Superconducting magnetic field detectors (SQUIDS) can detect neuronal currents without gross interference from tissue, but super cooling and large arrays are a problem.
- Detecting neuronal signals to muscles is not a problem but when such a signal exists, the muscle contracts and can be directly sensed.
- Brain signals are basically encrypted to us....

# Sensing Neuronal Activity

In addition to the direct detection of neuronal activity by directly inserting electrical probes into the brain, neuronal activity can be detected by less invasive techniques that sense electrical or magnetic fluctuations on the skin of the skull. The intervening body tissue is conductive and so shields and distorts electrical signals so that only very gross electrical brain activity can be observed. Magnetic fields only slightly interact with body tissue so in theory a very large array of magnetic detectors could be employed to map, in detail, electrical currents flowing in neurons. However today only SQUID (Superconducting Quantum Interference Devices) can detect such weak magnetic fields with fine spatial resolution. They require cooling to liquid-nitrogen temperatures and so must be insulated from the body. For the near future it is not likely that effective human-computer interfaces can employ these methods.

The ultimate problem with directly sensing brain signals is that there is so little known about how the brain functions that it is difficult to understand what the signals mean.

# The Visual Channel

- Raw resolution: ~1M pixels (from 100M rods & cones)
- 1/2 in fovea, 1/2 in surround
- Raw frequency response: ~ 20 Hz
- Limiting raw data rate is at optic nerve:
  - ~10Mb/s
  - Raw capacity 100 x greater other channels
- To match retina: ~ 1Gb/s
- To match & fill 360 degree field : ~$10^{13}$ b/s !
- Est. BW for a full resolution 'cave': ~$10^{9}$ b/s

# The Visual Channel

The human visual channel has, by two orders of magnitude, more raw bandwidth than other sensory channels. Thus it is the best target for improving the input interface to the human. Here we outline its properties.

# The High-Level Visual Channel

- Visual Comprehension is slower than seeing:
  - Speed reading text: ~ 300w/min = ~ 50b/s
  - Searching for keywords: ~200w/s = ~ 3000 b/s.
  - People recognize an icon faster than the word that stands for the icon. This is expected as there are far fewer icons that one knows than words.
- High performance in the visual channel will require specialized visual languages adapted to the task.
- Thus visual channel should employ multiple contexts each with a specialized visual language.
- Icons should probably be multidimensional
  - (3-d, animated, 'flickerable', dynamic colored & textured.)

# The High-Level Visual Channel

There is much room for improvement in the high-level visual channel. Reading text corresponds to only about 50 bits/sec. Higher performance should be possible by employing more effective visual languages. Icons of such a language should be multi-dimensional.

# Improving the Visual Channel

- Sensing gaze can allow a computer to adjust its display update rate and resolution as well as determine what objects other modalities are referencing.
- 1/2 of signal carried by optic nerve is not from the fovea. This resource is ignored by most of today's displays (an exception is the 'VR cave').
- The optic nerve carries short-time (flicker) information. Icons and other display objects can probably obtain additional visual bandwidth by employing 'flickering' or animated responses.

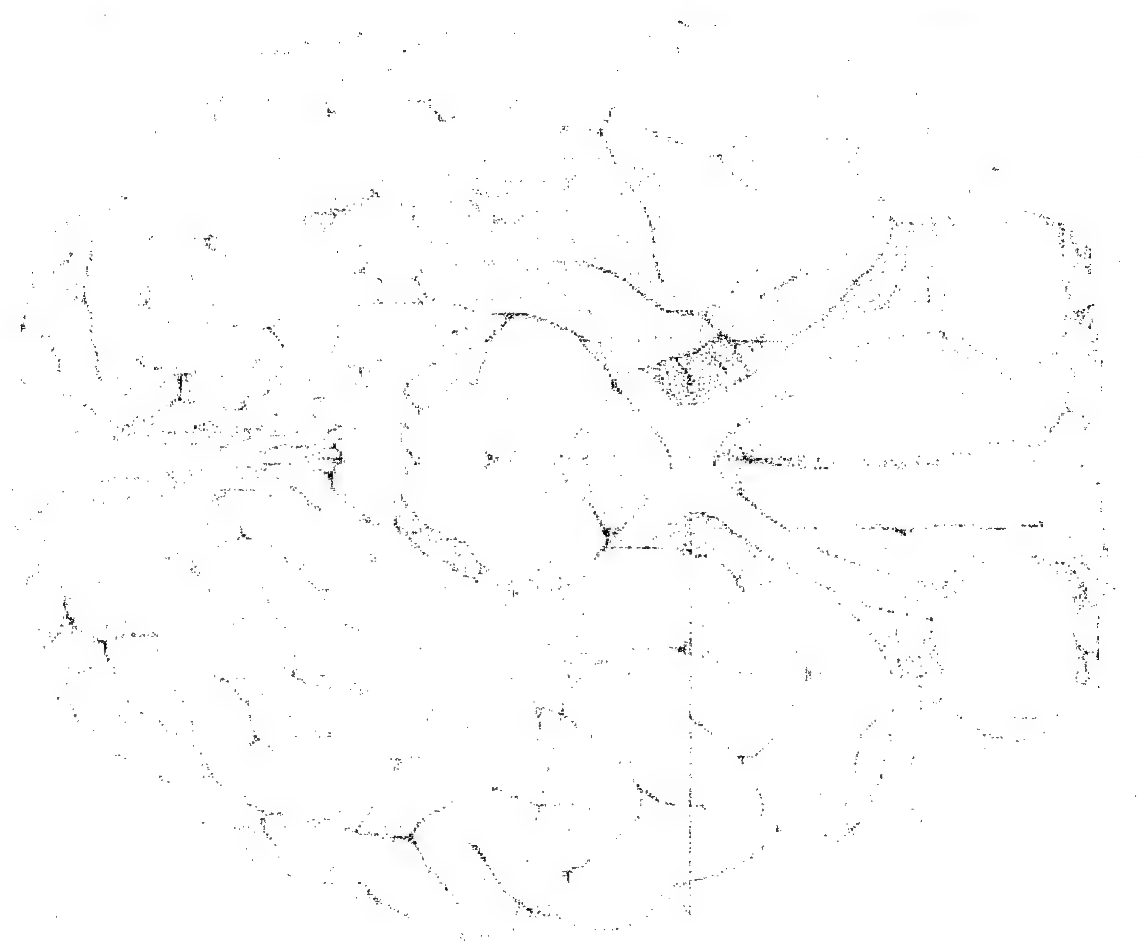
# Improving the Visual Channel

The visual channel can be enhanced by the computer tracking of gaze so resolution and update rates can be dynamically adapted to the visual channel.

The visual area outside the fovea is not generally employed in human-computer interfaces, yet about one-half of the optic nerve is dedicated to this area. Clearly this unused bandwidth could be employed to enhance performance.

The optic nerve also carries a great deal of information about short-time changes in light intensity. Additional information bandwidth in this channel could be utilized by employing ‘flickering’ and/or other animated features in visual displays.

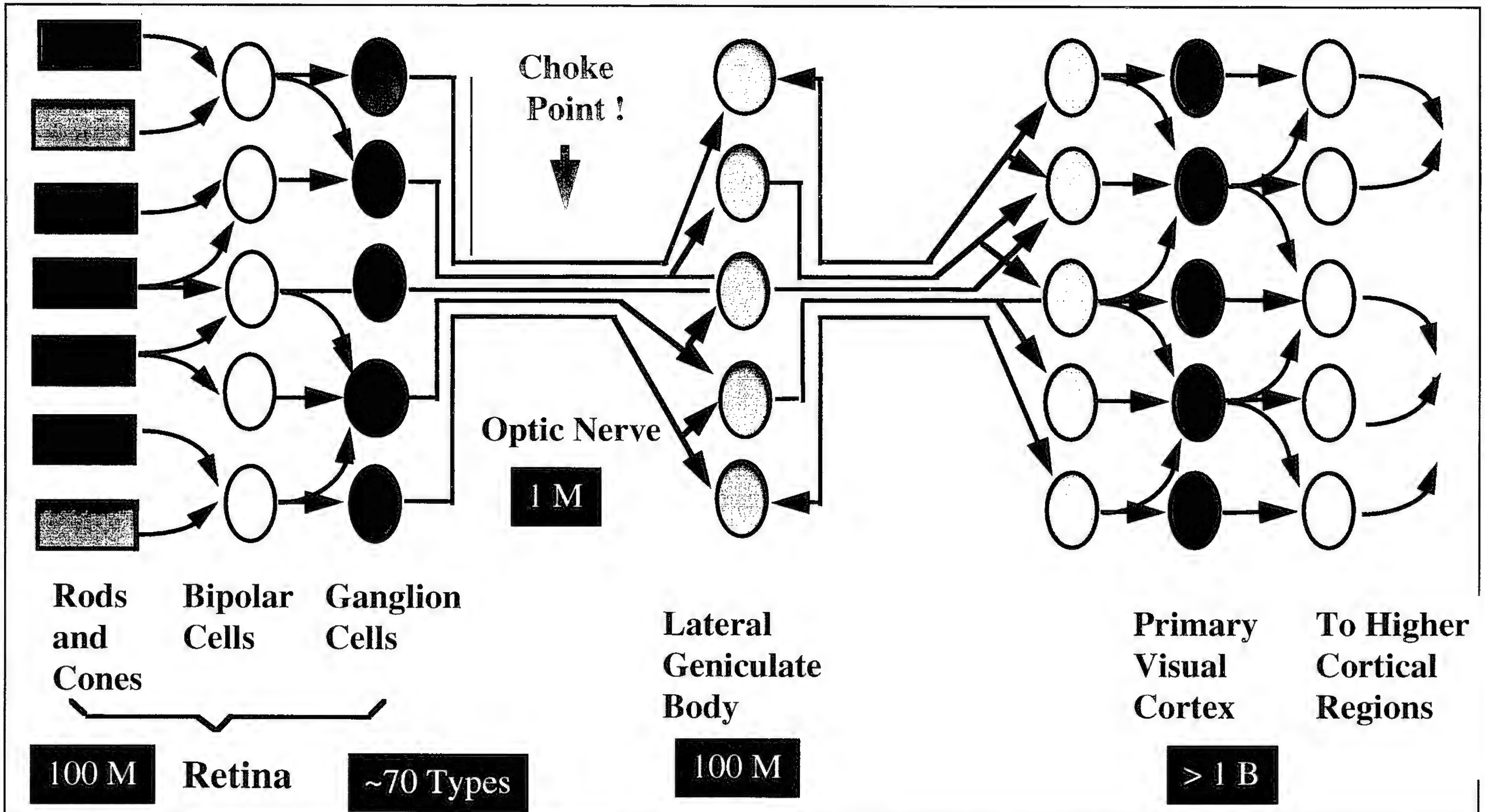
# Visual Cortex



# Visual Cortex

This is a drawing of a cross-section of the human brain illustrating the visual pathway from the eyes to the interbrain.

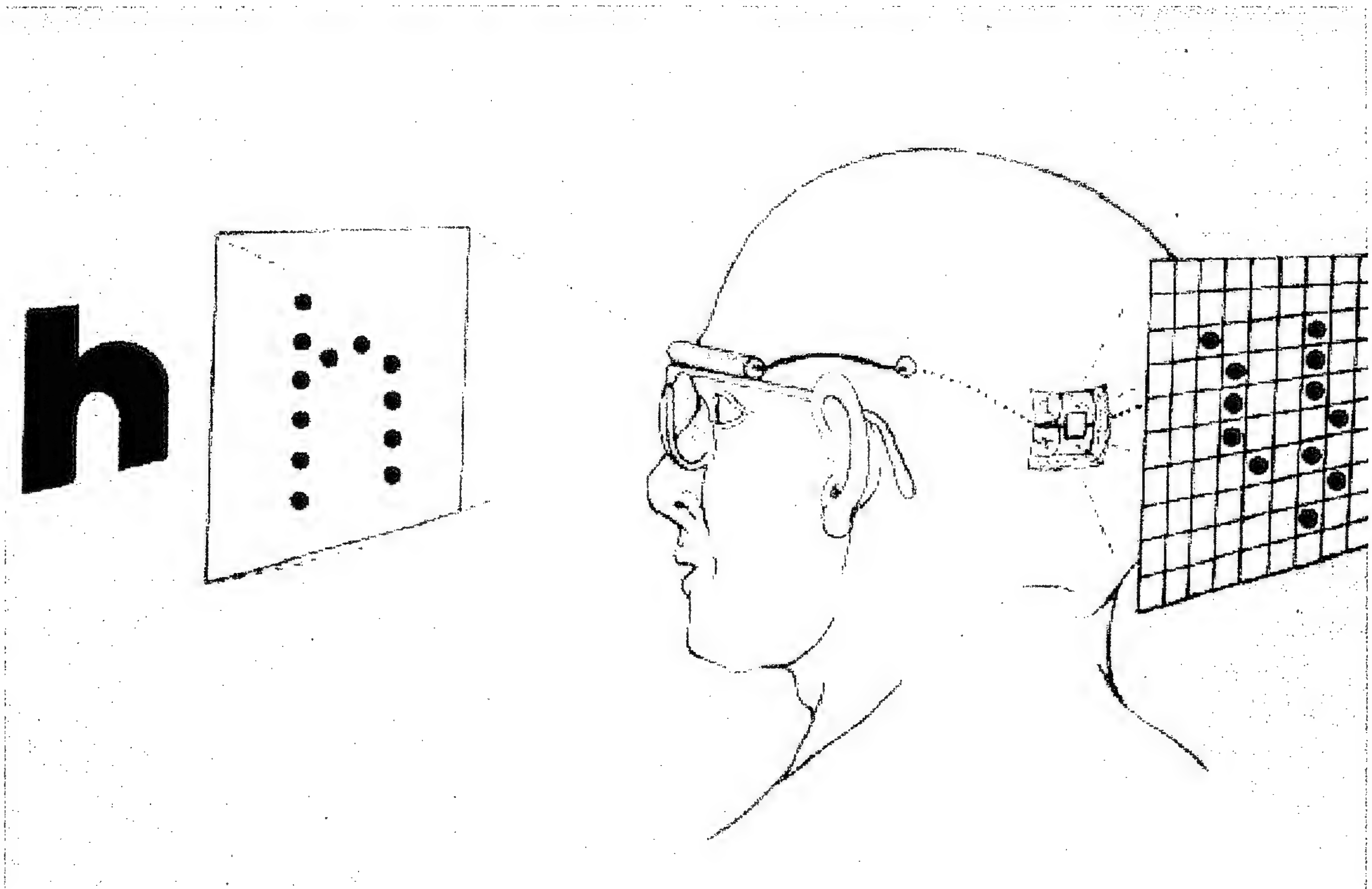
# Layers of the Visual Pathway



# Layers of the Visual Pathway

This is a highly simplified schematic representation of the visual pathway.

# Artificial Vision System



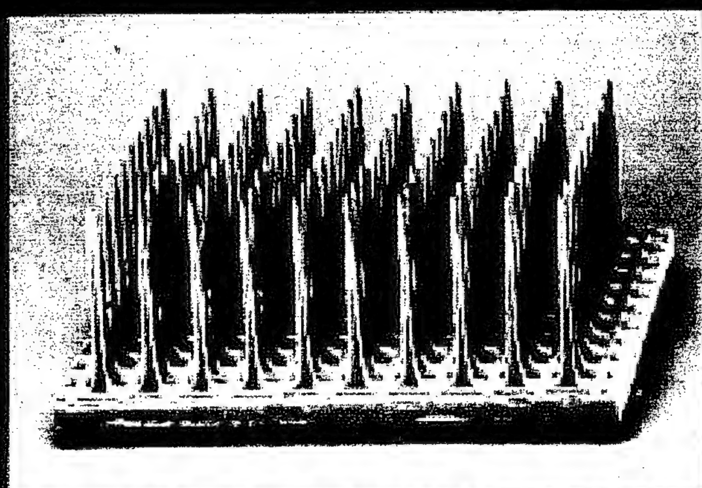
# Artificial Vision System

This is an illustration of a system being developed by Richard Normann's group at the University of Utah. The goal is to provide sight for the blind. Note the mapping from the visual field to the cortex field.

# Visual Cortex Implants

## The Utah Intracortical Electrode Array

- Cornerstone of the Artificial Vision Program -

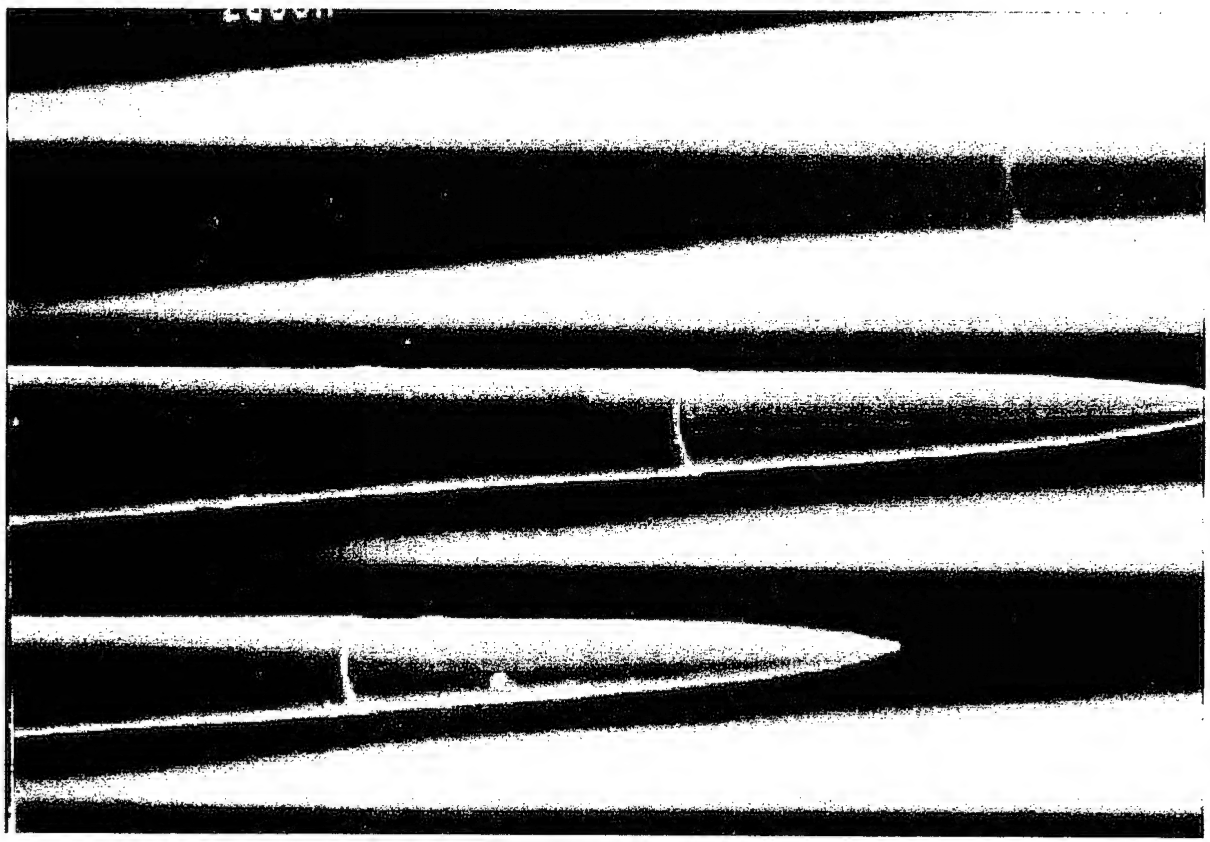


- ✓ Dimensions: 4.2mm x 4.2mm
- ✓ Electrode Length: 1.5mm
- ✓ Exposed Tip Length: ~.05mm
- ✓ Number of Electrodes: 100
- ✓ Electrodes are electrically isolated from neighbors

# Visual Cortex Implants

This illustrates the electrode array developed and employed by Normann's group at the University of Utah.

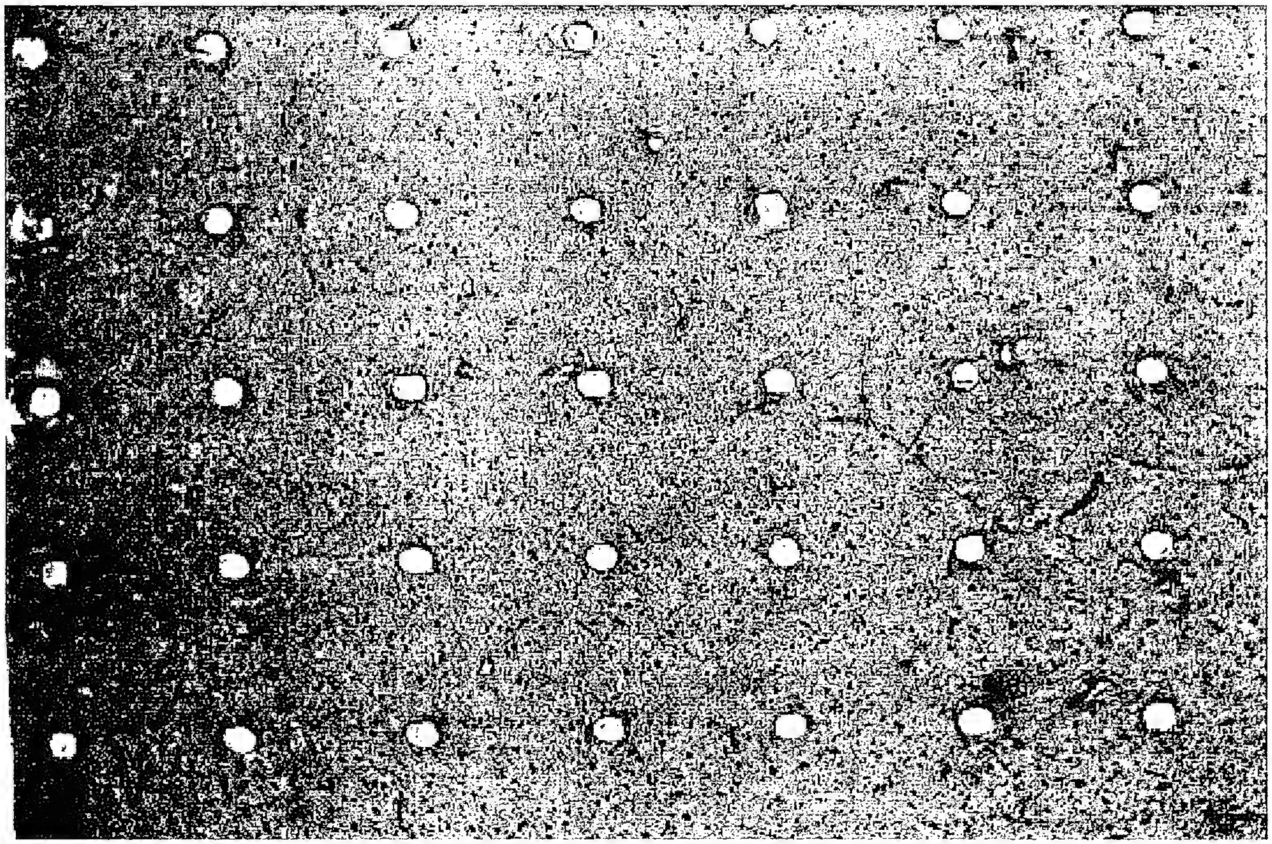
# Electrode Detail



# Electrode Detail

This is a close-up photograph (by the University of Utah) of the electrode array of the previous page. Note the insulated probes with the exposed tip.

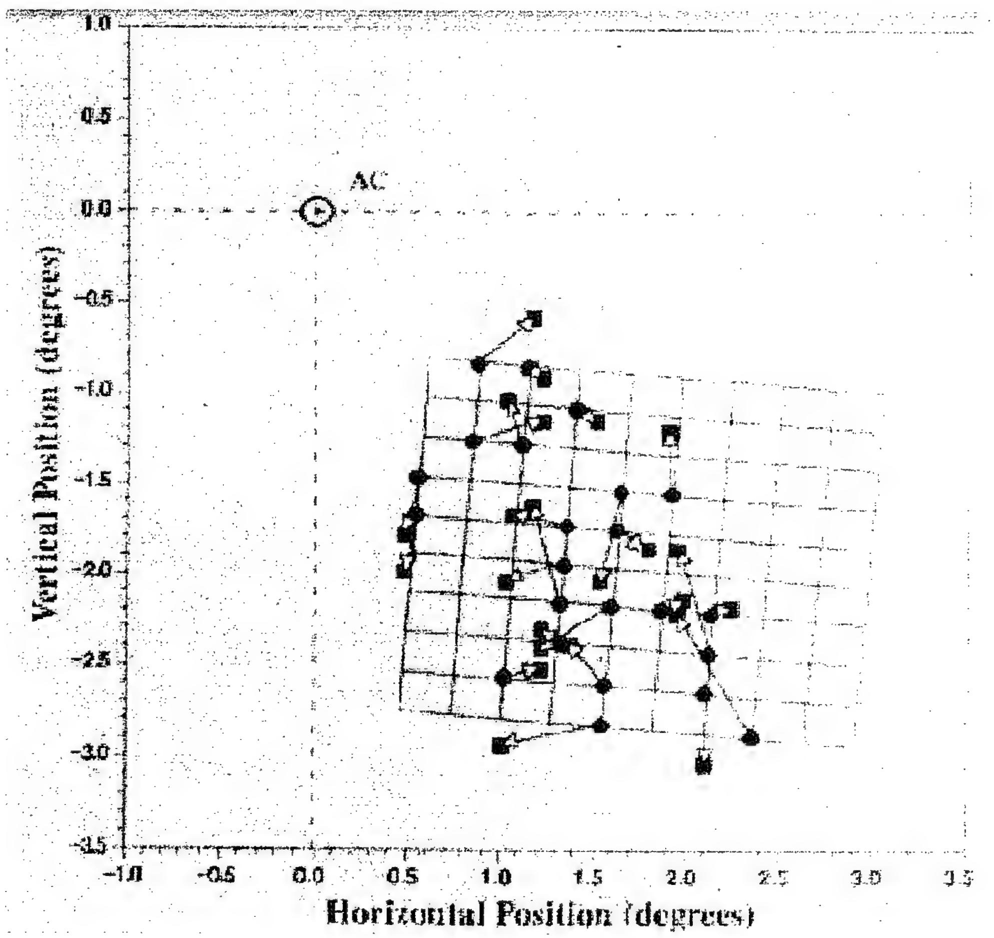
# Brain Tissue After Implant Removed



# Brain Tissue After Implant Removed

This is a photograph (by the University of Utah) of visual cortex after the probe array had been implanted for some time (several weeks) and then removed. What is important is that other than the physical holes, the brain tissue is normal, without inflammation, etc.

# Visual Map


AC
Vertical Position (degrees)
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
-3.0
-3.5
Horizontal Position (degrees)
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5

# Visual Map

This is an illustration of the visual mapping function, from the visual field to the cortex field, as measured on an experimental subject at the University of Utah. It is somewhat complicated but is systematic. The dots represent the electrodes, the squares the visual field. Corresponding pairs of dots and squares are connected by arrows.

# Conclusions re Direct Neural Input

- Active research is improving the connections
- Visual cortex implants (short term) have shown bandwidths greater than that of touch for reading Braille.
- However nothing like normal human visual acuity seems within reach for the next decade.
- Research may lead to a visual aid for the blind, but there is little possibility that any high bandwidth channel can be provided in the next few decades.
- The real problem is that very little understanding exists in detail as to how the visual system works.

# Conclusion re Direct Neural Input

While a direct interface to the visual cortex has been demonstrated, so far nothing like normal visual accuity has been achieved. The best result is an implant that a blind person used to visually read Braille faster than he could read the same text by touch.

This research work is important in reaching the goal of providing aid to the blind. However it does not offer much hope that an interface of higher performance than ordinary vision will be possible within the next decade. The problem is that there is too little understanding of how the brain works to design a high-performance interface even if much larger arrays of electrodes were to be used.

# The Human Auditory Channel

- Raw Sensitivity: 10Hz-40 kHz
- Raw data rate
  - CD sampling: 2 x 44k = 1.4 Mb/s
  - Telephone: 64 k b/s
- Speech
  - 10,000 word vocabulary: 14 bits
  - 3 words/s x 16 b/word = 40 b/s
- Music: 300 b/s
- A broad-band, poorly focused channel
  - It is much tougher to 'listen hard' than to stare
  - Good for background monitoring, warnings

# The Human Auditory Channel

At low levels of the auditory channel the bandwidth is about one megabit per second, far less than the lower levels of the visual channel. Interestingly, the bandwidth of the higher levels of speech and vision, processing computer generated information (typically text), are about the same at ~40 bits per second. Concentration for processing complex sound is more difficult to maintain for long periods however.

# Improving the Auditory Channel

- Various Modalities
  - Single tone, multi-voice, music
  - Pitch, volume, vibrato
  - Attack, sustain, harmonic content
  - Left/Right balance
- Questions arise:
  - What are the best modalities to encode artificial information?
  - In artificial spoken language what are the best phones to employ?
  - What is the interaction of the auditory with the visual channel?

# Improving the Auditory Channel

Cochlear implants are very successful for bringing some hearing to the deaf. They are good enough to allow the deaf to converse on a telephone. However it does not seem likely this will lead to any improvement over normal hearing for the same reasons as for visual implants: Brain function is not known in sufficient detail.

Improving the auditory channel for ordinary listening offers a number of possibilities including various modalities and various designs of audible languages.

# Improving Haptic (Direct contact) Interfaces

- Speed:
  - Muscles provide force
  - F=MA => smaller effective mass means faster response
  - Force-feedback reduces effective mass ('power steering')
- Completion:
  - Force-feedback 'taps' can indicate to fingers, the next expected series of key strokes.
  - If this 'feels right' user spaces, else continues.
  - Result is fewer strokes need be entered.
  - Performance improves.
- Vibration and temperature can also be used.
- Body has hundreds of muscles. More could be used. [face, body, legs,feet]

# Improving Haptic (Direct contact) Interfaces
(continued)

Haptic (direct physical contact) interfaces are the most commonly used input devices in the form of typewriter keyboards, and movement tracking devices - mice, trackballs, and trackpads. Haptic interfaces are less commonly used for computer output, aside from relatively simple applications - keys that click on completion of the keystroke, for example.

Haptic interfaces could be improved in a number of areas:

*Speed* - Muscles can provide force very rapidly, but the execution of physical motions requires more time because the mass of the hand, fingers and objects moved must be accelerated and decelerated. The inertial lag time can be reduced by reducing the motion required for input as for force sensitive "pencil eraser" input devices used in some laptop computers. When physical motion is desired, force feedback can be used to reduce the effective mass and time lag.

*Completion of motions* - Force feedback can also be used to suggest possible completions of actions for speed improvement and for training purposes. For example, a sequence of taps in a cording keyboard could be used to indicate the computer's guess at the completion of the word being typed. Force feedback could also be used to indicate a possible competion of a physical motion, by reducing the resistance along the computer's guides of the intended path.

# Improving Haptic (Direct contact) Interfaces

(concluded)

*Vibration, clicks, and temperature changes* - These additional physical attributes could be used to open new output channels from the computer to the user. For example, the frequency of vibration or temperature increase could be used to indicate proximity to some desired position.

*Body muscles* - Additional channels between computer and user could be opened by using more of the muscles in the body for input and output - facial expessions, head position, body position, legs and feet.

This page left blank intentionally.

# Input/Output via Physical Contact

- For humans:

  output rate ($\sim10^3$ bps) $\ll$ input rate ($\sim10^9$ bps)
  (hand motions) (vision)

- To maximize output use:
  - Mechanical devices to sense motion of:
    - All finger motion (~30 degrees of freedom) on both hands
    - Both hands, both arms, both feet, and body motions
  - Video camera to sense:
    - Eye gaze; winks, blinks, smiles, frowns, etc.
    - Hand and arm gestures

# Input/Output via Physical Contact

A fundamental problem for the human computer interface is the limited speed at which humans can output information via hand motions ~ $10^3$ bps. This data rate is far less than the rate ~ $10^9$ bps at which humans input data, primarily through vision.

To improve the bandwidth of the human/computer interface efforts should be made to sense the full range of human gestures and expressions. The fingers on each hand have ~ 20 degrees of freedom, and mechanical sensors could record all finger motions - grasping, pointing, curling of the fingers, etc. - not just keystrokes. One can imagine a variant of a chording keyboard in which the hands assume natural positions, and the characaters are coded in finger gestures. Such a device could be more comfortable and faster to use than a conventional keyboard, and could help users avoid the repetitive stress syndrome. Arm, feet, and body motions could also be used. A video camera trained on the user could be used to track eye gaze, and register facial expressions and gestures via pattern recognition software.

# Computer to Human Feedback

- Close coupled feedback from computer can greatly increase human output bandwidth
  - tactile sensing of virtual keys, objects, boundaries
  - guiding hand motions via mechanical feedback - anticipate probable actions
  - guiding eye gaze by flickering next gaze spot

# Computer to Human Feedback

Close coupled feedback from the computer can greatly increase the human output bandwidth.

Visual and tactile feedback could be used to create a virtual keyboard using two data gloves or a data muff (see below). Tactile feedback could be used to identify key positions and key clicks help identify key strokes. Data gloves or muffs could also be used to manipulate virtual objects, sense boundaries, and perform other virtual reality tasks.

Mechanical feedback can also be used to guide hand motions for training purposes, and to anticipate probable actions for faster response.

Eye gaze could be guided by using motion or flickering to attract attention, as done for advertisements on internet web pages.

# Force Sensing and Feedback

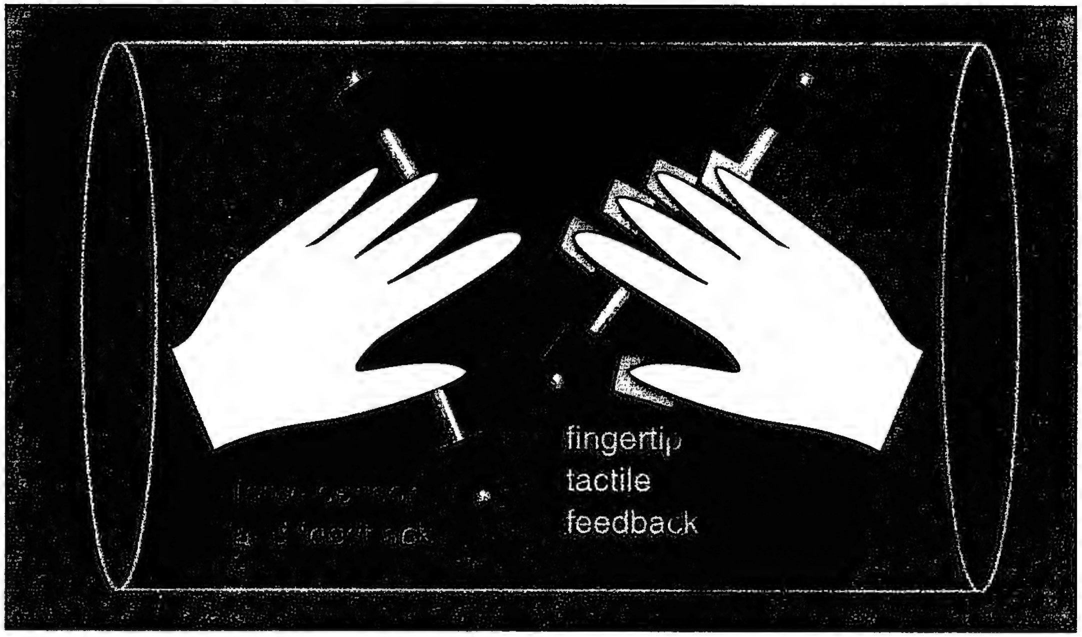

- Glove-based
  - strain sensor inputs
  - pneumatic outputs
- Force feedback used for
  - tactile sensing
  - "power steering" for the fingers
  - suggested motions

# Force Sensing and Feedback

Human hands are very sensitive and versatile and are likely to continue to be the primary interface for data input to computers.

A data glove equipped for force sensing and feedback is schematically illustrated in the figure. Force sensing could be implemented in the form of strain sensors incorporated into the fingers, and force feedback could be achieved via a series of pneumatic lines and bellows. The development of compact and flexible data gloves with high positional resolution and force sensitivity will continue to be an important goal for use in virtual reality systems. Data gloves with high resolution in position and force could be used for many interface tasks, ranging from the formation of a virtual keyboard (see above), to more sophisticated gesture recognition. For example the size and shape of an object being designed in a CAD system could be adjusted by simply grasping and manipulating the virtual object with the data glove. Many of the tricks currently used in CAD systems could be extended to this realm - using force feedback to snap the position to a grid, for example, or "magnetically" attracting objects to a virtual alignment bar.

# The Data Muff




- Sense forces with little hand motion - faster, not tiring
- Both force feedback and fingertip tactile feedback
- Engineering easier than glove-based systems.

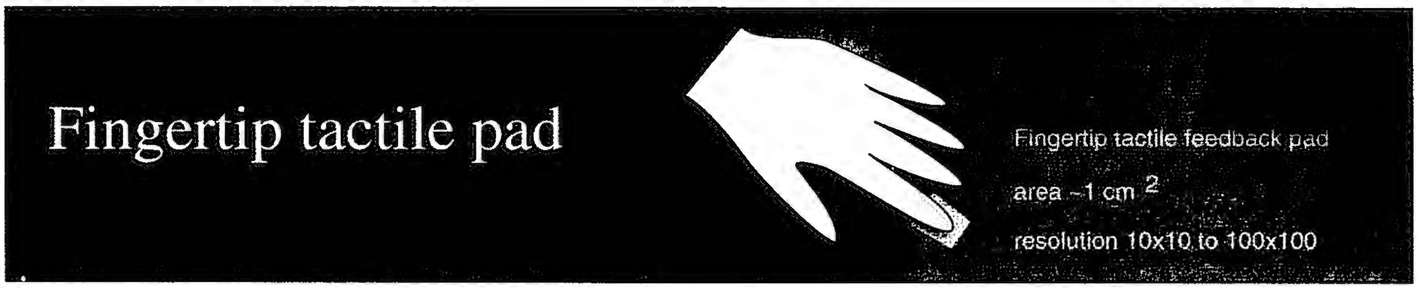
# The Data Muff

The data "muff" illustrated in the drawing may be a faster, higher performance version of two data gloves. The muff is fixed and contains two internal glove shaped cavities equipped with force sensors and feedback, as well as tactile sensing and feedback. Because forces are sensed and transmitted with little hand motion, the data muff should be faster and less tiring to use than data gloves. Because the muff is fixed, and need not be especially compact, the engineering design is much easier than for data gloves.

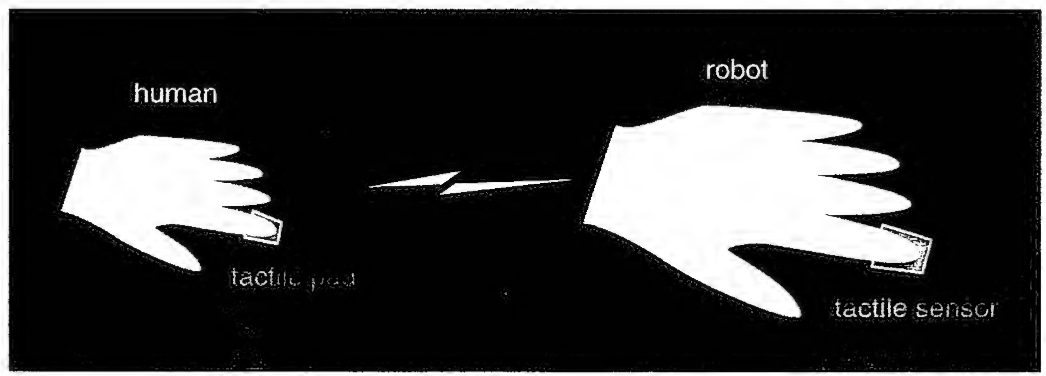
A data muff such as this could replace both the keyboard and the tracking devices of a conventional computer interface, as well as proved more advanced gesture recognition for CAD and other applications.

# Tactile Feedback

Fingertips very sensitive to motion (~1μm) and temperature; transients indicate contact, texture, wet/dry.




Uses:
- remote manipulators
- telepresence surgery
- virtual reality simulators

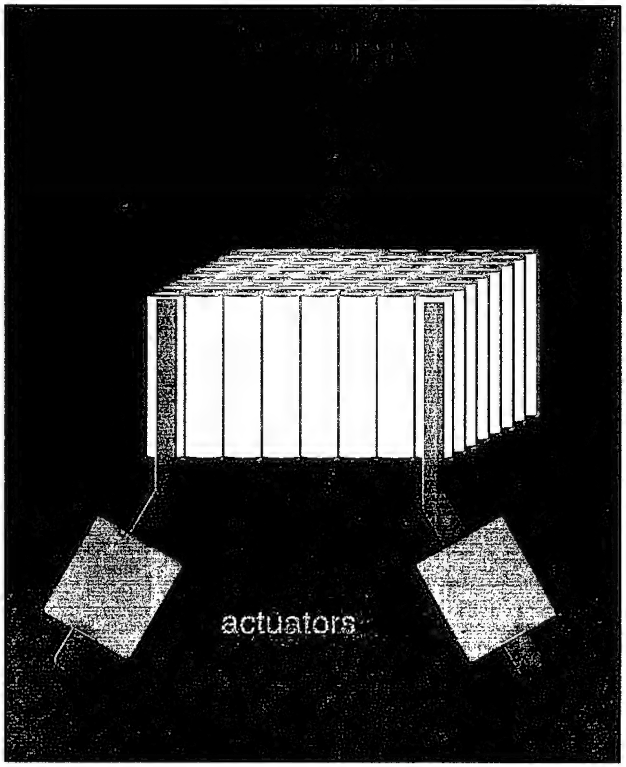


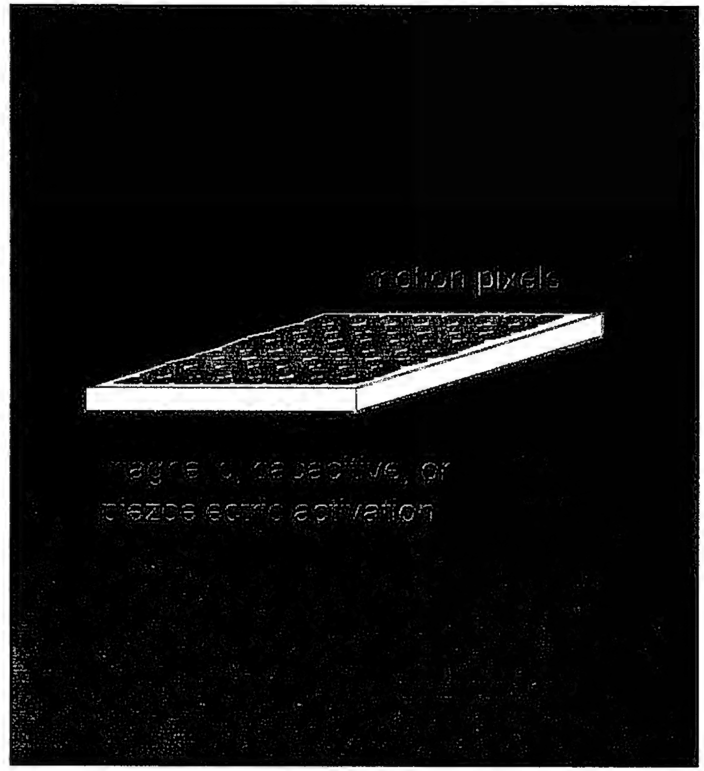
# Tactile Feedback

One area of haptic interfaces which has not been extensively developed is tactile feedback. Our fingertips are quite sensitive to motion (with ~ 1 μm resolution) and temperature. Transients indicate contact with an object, texture, and wet or dry conditions. These sensations convey a wealth of information which we use in everyday life, but have not been available in computer interfaces.

Using micro-electromechanical systems it should be possible to design and construct fingertip tactile feedback pads. The concept is illustrated by the figure: each pad could have an area - 1x1 $cm^2$ and express force and temperature with a spatial resolution from 10x10 pixels to 100x100 pixels. Feedback via tactile pads could have many uses. A set of tactile pads in a data glove or muff mated with a corresponding set on a remotely located robotic hand could provide the sensitivity for tasks which are currently very difficult. For example, the operator could use the robotic hand to reach into a jar of nuts and bolts, pick out a 6–32 nut by feel, and thread it onto a machine screw. Telepresence surgery is an important application of robotics currently under development, where the addition of tactile feedback could be of benefit - for grasping tools and sensing the condition of the patient. Virtual reality simulators could provide a wide range of applications for tactile feedback.

# Fingertip Pad



# Fingertip Pad

Two approaches to tactile feedback of motion and textures are illustrated, one based on an array of actuated pins, and a more advanced concept using micro electromechanical systems ideas. Thermal feedback with spatial resolution could be achieved via an array of small electrical heaters, one on each pixel.

Mechanically actuated pin arrays analogous to printer heads are currently used to output Braille characters to blind users. The technology is already developed and works well. The disadvantages of this approach are relatively large size and weight, complexity of manufacture, and limited spatial resolution.

Micro-electromechanical techniques could be used to make an integrated fingertip tactile pad with potentially higher spatial resolution and lower weight. Each pixel in an array could be mechanically actuated via magnetic (current in the presence of a magnetic field), capacitive (attraction of charged capacitor plates) forces. Piezoelectric actuators could also be used to produce textures by amplifying the small motion of the piezoelectric pixels with mechanical bubbles or cantilevers.

# Unobtrusive Interfaces

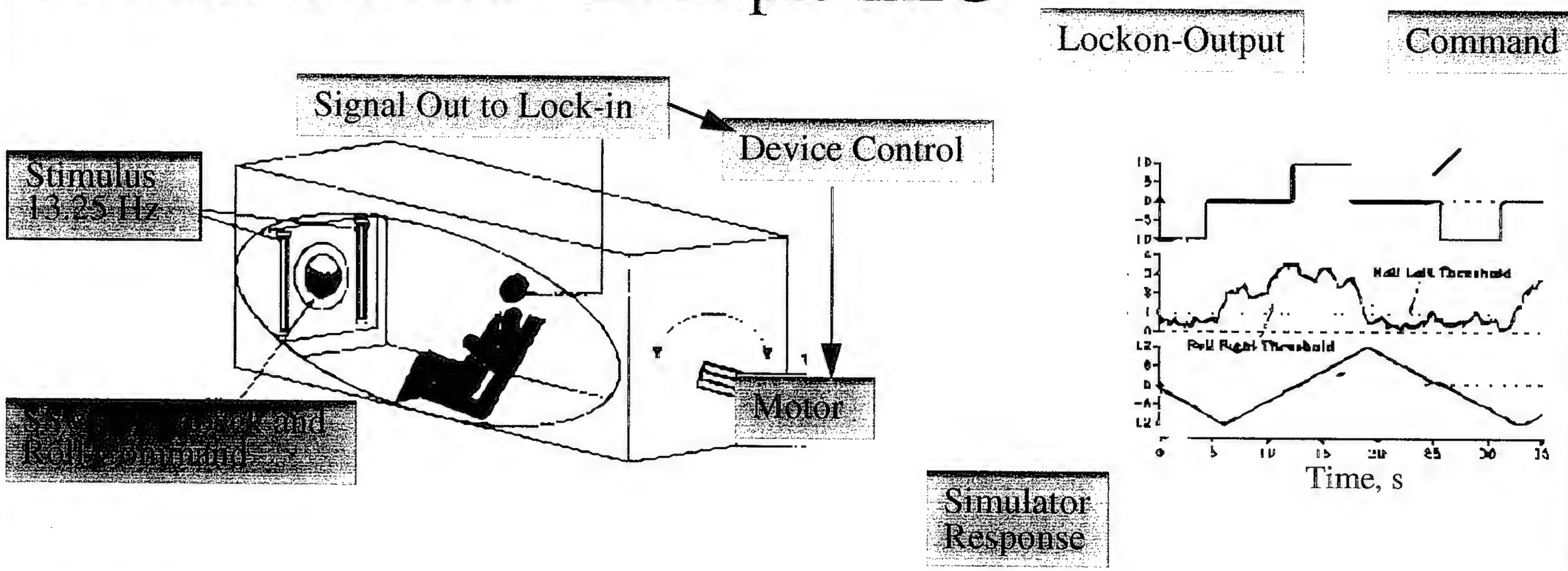
- Goal: Improve human efficiency by reducing conscious attention required to operate computer
- Approach: Computer sensors monitor human intention, respond to human command
- Context: Computer as "cognitive tool"
- Hardware and Software Possibilities
  - Eye Tracker + high level location identification
  - Screen superposition display keyed to eye direction
  - Video camera + facial and motion recognition
  - Audio Detector + Low Level speech recognition
  - Other sensors (e.g. tactile, EEG, EM.....)

# Unobtrusive Interfaces

Unobtrusive interfaces operate without conscious attention from the user. This leaves the user free to concentrate on the task at hand. The basic idea is to instrument the user work station to measure a wide variety of human activities. For example eye gaze, facial expression, head position, body motion, human sounds (not limited to speech), muscle tension and nerve activity could all be sensed. The computer then adapts to these inputs to improve the interface. Several examples follow.

# EEG Interface

- Other Sensors - Example EEG




Biofeedback requires extensive (10's of hours) training for good performance (80-95% success with single-axis task). Response times of several seconds have been achieved.

G.R. McMillan and G.L. Calhoun, RESNA '95

# EEG Interface

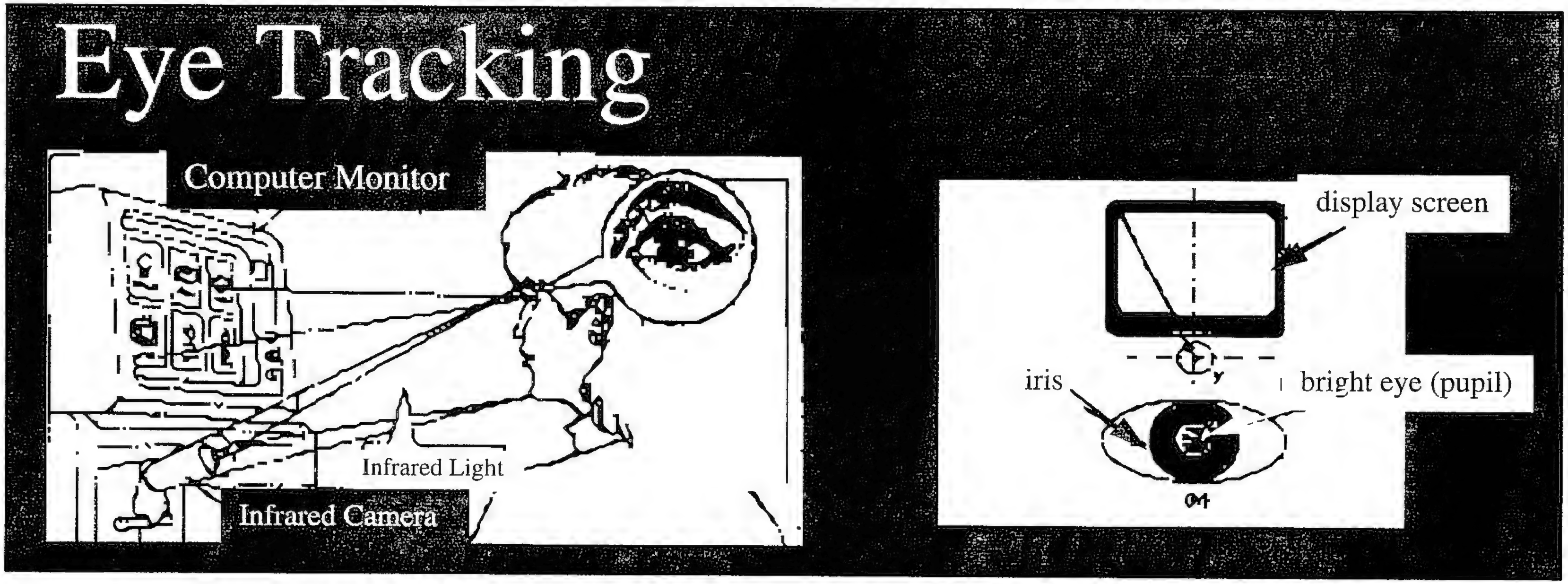
The U. S. Air Force and others have tried for many years to control machines (e.g. airplanes) by thought. A recent experiment, illustrated above, demonstrate this idea. The experiment simulates a very simple flight trainer. There is only one axis of control. The idea is to determine how well an operator can control roll.

The operator views an artificial horizon and attempts to keep his “vehicle” level in the face of disturbances, etc. On the sides of the display there are light panels, whose intensity is modulated by a 13.25 Hz signal. The peripheral vision system process this signal and the result is detected by EEG and a lock-in amplifier.

With several ten’s of hours of training the operator can learn to modulate the EEG signal enough to control the single axis motor. The response times of several seconds is very slow however. The bandwidth of this channel is thus only about one bit per second.

While this is an interesting experiment, this approach (sensing brain activity using EEG) seems to offer little toward the goal of a high performance interface.

# Eye Gaze Interface




Low-power near-infrared source and camera provide eye look-point to ≈ 0.3 °
by comparison of corneal reflection (glint) and image of pupil center

Eye speed between points of gaze essentially infinite (400-600 °/s)

Independent measure of head position required for externally mounted tracker

# Eye Gaze Interface

There is a well-developed technology for tracking the gaze of the eye. Accuracy is about 0.3 degrees and the speed is about 500 degrees per second. For a work station, such an eye tracking system could replace the mouse employing an eye wink for the mouse for curser positioning by button activation (blinks would be ignored.) This could considerably speed-up the interface to workstations.

# Summary - Unobtrusive Interface

- Implementation
  - Clicking functions (locate, activate, undo) implemented via nod, eye motion to icon, voice command, hand motion, other...
  - Head position via magnetic sensors ~ 0.1 x 0.1degrees
  - Computer indication (highlight) of stable gaze points (0.1-0.2s) for user feedback
  - Voice control limited to simple commands to avoid slow response ( 1 s for 2-5 word strings) of speech recognition
  - Perfect resolution display via superposition of full screen and demagnified high resolution image of gaze point region ( $1^{\circ}$ FOV)
  - All components are demonstrated in commercial or research instrumentation: unobtrusive interface requires "good-enough" combination of options

# Summary -Unobtrusive Interfaces

Unobtrusive interfaces have the potential to significantly improve the performance of the human-computer interface. Video cameras can determine head position, body-language, and gestures. IR scanners can determine gaze. Voice recognition and interpretation of other sounds (clicks, etc.) can also help. Foot pedals and other body sensors are also useful. Vibrators, speakers, peripheral displays, seat shakers and direct electrical stimulation can all increase the performance of the computer to human channel.

In summary, there are a number of sensors, activators, and processing techniques available to improve the human-computer interface. So far, only a few isolated experiments have tried more than one modality at a time. What is needed is a systematic series of experiments where all these are integrated to work together synergistically. This may be a unique opportunity for DARPA.

# Levels of Abstraction - User Models

- [n-D]: Group Coordination&Action/Web Activity
- [3-D]: Graphics, Data gloves, Caves
- [2-D]: Word processor, Spread Sheets, Mice
- [1-D]: Command Line, Linear printing
- [0-D]: Front panel switches and indicator lights

# Levels of Abstraction - User Models

User models can be categorized into an abstraction hierarchy corresponding to geometric dimensions. Early computer systems employed lower levels while modern PC's and workstation today typically use two dimensional user interfaces. Previously in this presentation we have considered all three geometrical dimensions. We will now examine issues that employ additional dimensions (at higher levels of abstraction.) For example a software agent can act on behalf of a user at distant times and places.

# Digital Personal Communication Assistants

- Like a spouse, butler, or elite team member your Digital Personal Communication Assistant should be able to complete before you do your:
  - finger movements toward keys
  - words to be typed or spoken
  - phrases, sentences, commands, requests
- It should be able to anticipate what data, messages, documents, searches and communication sessions you probably want next (or soon) and prefetch them so you don't generally have to wait for information.

# Digital Personal Communication Assistants

(concluded)

- Main functions include:
  - Maintaining a cache of predicted useful materials
  - parsing the language of the interface
  - pre-fetching information
  - formating to the user's taste
  - adaptation to the user and his current activities
- Such Assistants, each different, are needed at each level of the communication hierarchy
- Such Assistants are much easier to achieve than general artificial intelligence
  - No deep reasoning or extensive knowledge is required.

# Digital Personal Communication Assistants

One simple form of a software agent is a software personal assistant agent. Here we examine such assistants specialized to help with communications. There are many functions that such assistants can help with. Some important ones are listed here.

Thad Starner (Cyborg) employs a text editor that anticipates what word he is starting to type on his single hand 'twiddler' keyboard. This word is predicted based on his past typing history and a on-line dictionary. After each keystroke, the predicted following letters are displayed. If the predicted word is correct he spaces and starts the next word, otherwise he keeps typing as usual. Starner reports that he rarely types more than three characters per word. This allows him to record all conversations people have with him as they occur and with no noticeable interference.

Another possible Digital Personal Communication Assistant (A Web Assistant) could help a user more effectively locate and employ information on the world wide web. For example a Web Assistant might learn to prefetch that information on the web that a user regularity reads, e.g. the Wall Street Week web edition. Thus the user would not need to wait for a slow down-load.

This page left blank intentionally.

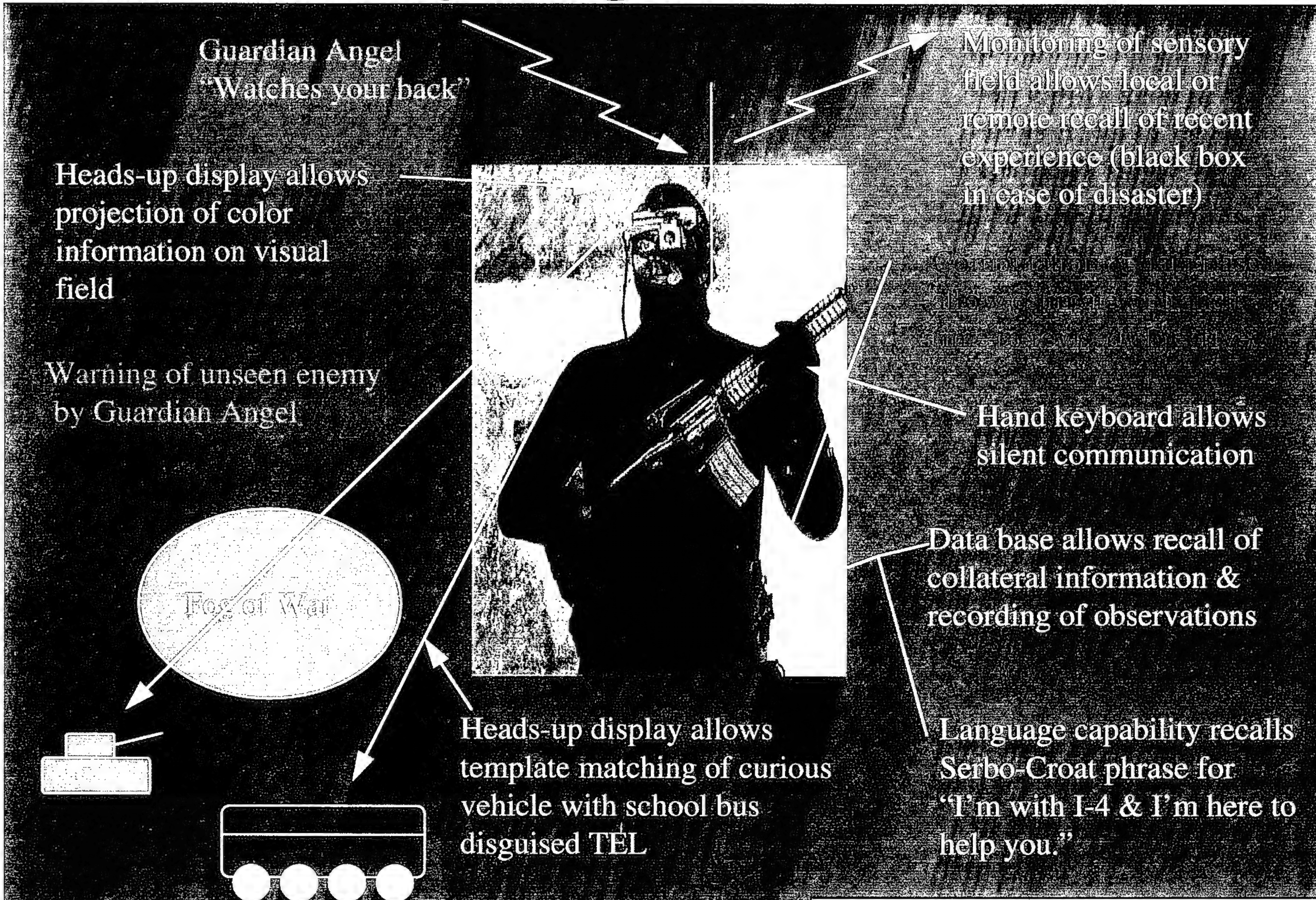
# Networked Systems

- Includes groupware, web conversation rooms, collaborative design, enterprise integration
- This JASON study was not able to devote much time to this very important topic
- Clearly there is much activity in this area and it is an area with great possibilities
- Commercial development will continue for ordinary users
- Best opportunity for DARPA may be in the support for elite teams (Seals, etc.)

# Networked Systems

The JASON Summer Study time period ended before the very interesting topic of networked human-computer interfaces could be investigated. However. it seems to be an area rich in possibilities and with potential pay-offs for supporting elite teams.

# Cyborg Warrior



# Cyborg Warrior

The Cyborg Warrior incorporates many of the innovations enabled by advanced computation & wireless network communications. Working counterclockwise from the top note the `Guardian Angel' that keeps track of possible dangers to the warrior and warns him at critical times. Such warnings and other pertinent or requested information can be transmitted by means of a heads-up display. The warrior's visual field could include infrared night vision as well as enhanced visual optics. The 'heads-up' display could supply templates to make detailed differentiation between targets, decoys, non-targets and disguised targets. For operation in foreign lands, the cyborg warrior can call upon language phrases to converse with foreign soldiers and civilians.

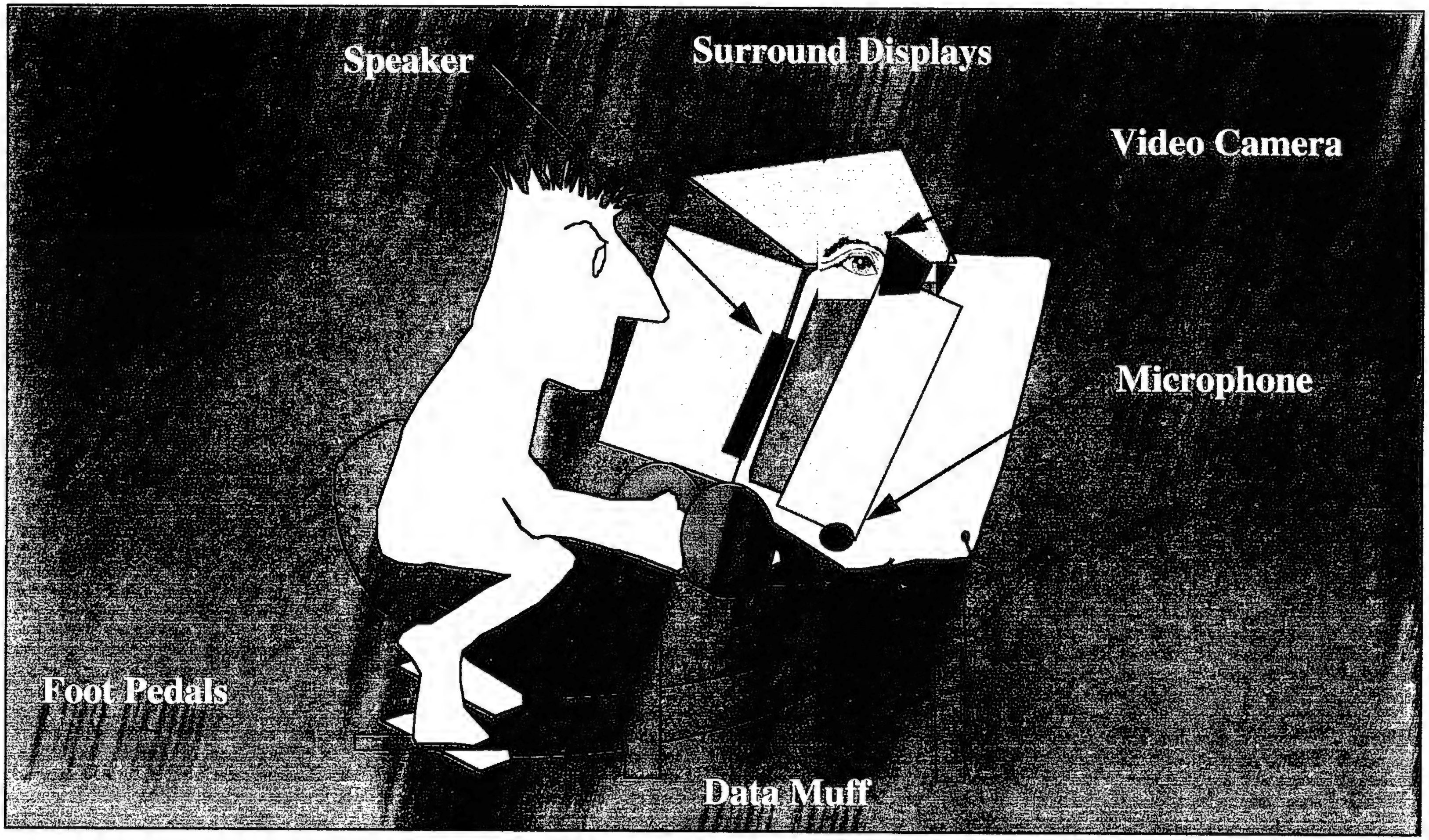
# Cyborg Experiments as Testbed of Human/Computer Relationship

- Cyborgs teams can have highly coordinated and externally supported information systems
- Cyborg experiments allow assessment of practical aspects of intimate human/computer relationship
- Durability and robust character of essential equipment, e.g. CPU, disc drive, display, etc.
- Information display optimization -- how much, text or graphics, control of display, usefulness of color, etc.
- Long term, real world tests of interfaces such as "Twiddler" type one-handed keyboard

# Cyborg Experiments as Testbed of Human/Computer Relationship

A resident data base allows the warrior to consult pertinent information from other sources as well as data he has entered himself using a one-hand keyboard. This keyboard also allows silent communication. Combination of entered data and geographical information allows progress toward an objective to be recorded and a time of arrival estimated. The data base can also act as a `black box' recording sensor and other information for retrievals. This would be especially useful if the warrior were injured and unable to communicate in the normal way. The warrior's health status alone could indicate that help should be sent and to which location. The Cyborg warrior gets the advantage over his opponents by enhanced intelligence, both received and transmitted. He can achieve the Duke of Wellington's wish to be able to 'see over the hill.'

# Professional Workstation



# Professional Workstation

A new generation of professional workstations could benefit from the types of human-computer interfaces discussed above. Here we illustrate how some of these might be integrated into a professional workstation. The visual channel is enhanced by the use of surround displays. A video camera and infra-red scanner convey the position of the user's gaze and his facial dynamics (winks, etc) to the computer. Microphone, foot pedals and the 'Data Muff' also enhance the user's ability to input data to the computer. Not shown are the several 'Digital Personal Communication Assistants' (software agents.)

# Conclusions

- The highest performance will be achieved by :
  - Identifying and characterizing each level of the human hierarchy
  - Given the five human senses and one motor channel, the roughly six layers of sensor and motor neuron layers, this implies the hierarchy is ~10 levels deep and has at least 30 separately organized mechanisms.
- Raw bandwidth can possibly be increased by physical and biological means such as drugs, etc.
- Better and more haptic interfaces can be employed as can better visual displays.

# Conclusions

(continued)

- Feedback channels can add to performance.
  - Computer communication processes that model the corresponding human levels can improve performance.
- There is a lot of 'missing science' needed to make progress in Human-Computer Interfaces (HCI):
  - What the 'unknown 90%' types of neurons in the retina do
  - What the models are for each level of the human hierarchy
  - How the brain processes visual images (in detail)

# Conclusions

(concluded)

- A computer model that attempts to model the human communication channel hierarchy as well as we know it would help not only in research to improve understanding but in the design of future HCI's.
- New physical interface devices could be developed to improve channel bandwidth for professionals

# Conclusions

In conclusion it can be seen that there are many possibilities for achieving a high performance human-computer interface for professionals, both in the field and in the office. Research is needed to identify and characterize all the levels of human communication hierarchy. Then computer communication processes can be designed to match the human mechanisms. Feedback and stored knowledge will be important at each level. Better haptic interfaces can offer enhanced performance as well.

It is striking how little is understood about how humans process complex information. There is a lot of 'missing science' that will need to be developed to fully exploit all the possibilities for improving the human-computer interface. Because the visual channel has such high bandwidth, the greatest need is to better understand human visual processing.

# Some Possible Projects

- Go after the 'missing science'
- Build and test an 'anthropomorphic keyboard':
  - Non-rectangular key layout
  - force-feedback keys
  - dynamic adaptation of key force to optimize speed
  - put 'mouse functions' into the keys (via force on key)
  - Use more aggressive GUI (harder to learn, but faster)
- Try building the 'Data Muff'
- Try building the 'Professional Workstation'
- Try improving each of the extant Cyborg components (they need work)

# Some Possible Projects

There are a large number of opportunities for DARPA to develop higher performance human-concept interfaces. DARPA should not follow the industry trend of developing simpler interfaces for a growing public of computer users, but should concentrate on the area neglected by today's mass market computer manufacturers -- the high performance (even if difficult to learn) interfaces. Projects in fundamental science, in new experimental haptic devices, and in integrated interfaces such as the 'Professional Workstation' and 'Elite team Cyborg Interface' could all contribute to improved high-performance human-computer interfaces.

## DISTRIBUTION LIST

DARPA Library
3701 North Fairfax Drive
Arlington, VA 22209-2308

DTIC [2]
8725 John Jay Kingman Road
Suite 0944
Fort Belvoir, VA 22060-6218

Director of Space and SDI Programs
SAF/AQSC
1060 Air Force Pentagon
Washington, DC 20330-1060

Superintendent
Code 1424
Attn Documents Librarian
Naval Postgraduate School
Monterey, CA 93943

CMDR & Program Executive Officer
U S Army/CSSD-ZA
Strategic Defense Command
PO Box 15280
Arlington, VA 22215-0150

Director
Technology Directorate
Office of Naval Research
Room 407
800 N. Quincy Street
Arlington, VA 20305-1000

Dr Albert Brandenstein
Chief Scientist
Office of Nat'l Drug Control Policy
Executive Office of the President
Washington, DC 20500

Dr H Lee Buchanan, I I I
Director
ARPA/DSO
3701 North Fairfax Drive
Arlington, VA 22203-1714

Dr Collier
Chief Scientist
U S Army Strategic Defense Command
PO Box 15280
Arlington, VA 22215-0280

Dr Victor Demarines, Jr.
President and Chief Exec Officer
The MITRE Corporation
A20
202 Burlington Road
A210
Bedfored, MA 01730-1420

Mr Dan Flynn [5]
OSWR
Washington, DC 20505

Dr Paris Genalis
Deputy Director
OUSD(A&T)/S&TS/NW
The Pentagon, Room 3D1048
Washington, DC 20301

Dr Lawrence K. Gershwin
NIC/NIO/S&T
7E47, OHB
Washington, DC 20505

Dr Robert G Henderson
Director
JASON Program Office
The MITRE Corporation
7525 Colshire Drive
Mailstop Z561
McLean, VA 22102

## DISTRIBUTION LIST

Dr William E Howard I I I [2]
Director of Advanced Concepts &
Systems Design
The Pentagon Room 3E480
Washington, DC 20301-0103

J A S O N Library [5]
The MITRE Corporation
Mail Stop W002
7525 Colshire Drive
McLean, VA 22102

Dr Anita Jones
Department of Defense
DOD, DDR&E
The Pentagon, Room 3E1014
Washington, DC 20301

Mr. O' Dean P. Judd
Los Alamos National Laboratory
Mailstop F650
Los Alamos, NM 87545

Dr Bobby R Junker
Office of Naval Research
Code 111
800 North Quincy Street
Arlington, VA 22217

Dr Ken Kress
Office of Research and Development
809 Ames Building
Washington, DC 20505

Lt Gen, Howard W. Leaf, ( Retired)
Director, Test and Evaluation
HQ USAF/TE
1650 Air Force Pentagon
Washington, DC 20330-1650

Mr. Larry Lynn
Director
DARPA/DIRO
3701 North Fairfax Drive
Arlington, VA 22203-1714

Dr. John Lyons
Director of Corporate Laboratory
US Army Laboratory Command
2800 Powder Mill Road
Adelphi,MD 20783-1145

Col Ed Mahen
ARPA/DIRO
3701 North Fairfax Drive
Arlington, VA 22203-1714

Dr. Arthur Manfredi
OSWR
Washington, DC 20505

Mr James J Mattice
Deputy Asst Secretary
(Research & Engineering), SAF/AQ
Pentagon, Room 4D-977
Washington, DC 20330-1000

Dr George Mayer
Office of Director of Defense
Reserach and Engineering
Pentagon, Room 3D375
Washington, DC 20301-3030

Dr Bill Murphy
ORD
Washington, DC 20505

Dr Julian C Nall
Institute for Defense Analyses
1801 North Beauregard Street
Alexandria, VA 22311

## DISTRIBUTION LIST

Dr Ari Patrinos [5]
Director
Environmental Sciences Division
ER74/GTN
US Department of Energy
Washington, DC 20585

Dr Bruce Pierce
USD(A)D S
The Pentagon, Room 3D136
Washington, DC 20301-3090

Mr John Rausch [2]
Division Head 06 Department
NAVOPINTCEN
4301 Suitland Road
Washington, DC 20390

Records Resource
The MITRE Corporation
Mailstop W115
7525 Colshire Drive
McLean,VA 22102

Dr Victor H Reis [5]
US Department of Energy
DP-1, Room 4A019
1000 Independence Ave, SW
Washington, DC 20585

Dr Fred E Saalfeld
Director
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

Dr Dan Schuresko
O/DDS&T
Washington, DC 20505

Dr John Schuster
Technical Director of Submarine
and SSBN Security Program
Department of the Navy OP-02T
The Pentagon Room 4D534
Washington, DC 20350-2000

Dr Michael A Stroscio
US Army Research Office
P. O. Box 12211
Research Triangle NC 27709-2211

Ambassador James Sweeney
Chief Science Advisor
USACDA
320 21st Street NW
Washington, DC 20451

Dr George W Ullrich [3]
Deputy Director
Defense Nuclear Agency
6801 Telegraph Road
Alexandria, VA 22310

Dr David Whelan
DARPA/TTO
3701 North Fairfax Drive
Arlington, VA 2220-1714

Dr Edward C Whitman
Dep Assistant Secretary of the Navy
C3I Electronic Warfare & Space
Department of the Navy
The Pentagon 4D745
Washington, DC 20350-5000